

The University of Chicago

ELEMENTS OF READING MATERIALS
CONTRIBUTING TO DIFFICULTIES
IN COMPREHENSION ON THE
PART OF ADULTS

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF
THE SOCIAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

1933

By

BERNICE ELIZABETH LEARY

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1935

The University of Chicago

ELEMENTS OF READING MATERIALS
CONTRIBUTING TO DIFFICULTIES
IN COMPREHENSION ON THE
PART OF ADULTS

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF
THE SOCIAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

1933

By

BERNICE ELIZABETH LEARY

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1935

TABLE OF CONTENTS

	Page
List of Tables	111
The Problem	1
Defining Potential Elements of Difficulty	2
Selecting Reading Materials	8
Constructing Adult Reading Tests	13
Identifying Elements That Influence Difficulty	17
Estimating Difficulty of Reading Materials	25
Objectives of Future Investigation	33
Bibliography	34



LIST OF TABLES

Table	Page
I. Expressional Variants Suggested as Potential Elements of Difficulty in Adult Reading Material . . .	3
II. Names of Twenty-nine Books Selected as Sources of Test Items, with Classification and Author of Each.	9
III. Magazines Selected as Sources of Test Items	10
IV. Names of Fifteen Newspapers Selected as Sources of Test Items.	11
V. Classification of Reading Material into Groups of Probable Difficulty	12
VI. Name, Race, Classification, Grade Placement, Geographical Location, and Size of Twenty-one Groups of Adults Tested by Both Form 1 and Form 2 of the Adult Reading Test.	19
VII. Criteria of Difficulty for Items on Form 1, Represented by the Average Score of All Readers, "Best" Readers, and "Poorest" Readers.	21
VIII. Criteria of Difficulty for Items on Form 2, Represented by the Average Score of All Readers, "Best" Readers, and "Poorest" Readers.	21
IX. Correlations between Forty-four Potential Elements of Difficulty and the Average Scores of Each Test Item for Three Classifications of Adult Subjects. .	23
X. Correlations of Significant Elements of Difficulty with Average Reading Score and with Each Other. . .	27
XI. Instruments of Prediction, Combining Four Elements of Difficulty	31



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548807_0104

The Problem

An outstanding problem in administering a program of adult education is the provision of reading materials to meet the needs of adults with mature interests and limited reading abilities. Although writers and publishers have attempted to adapt books to these needs by writing in a simplified language, evidence has accumulated from librarians and teachers of adult classes which indicates that too many books are written in terminology quite incomprehensible to readers of low ability.¹ Furthermore, librarians maintain that they have no reliable technique by means of which to estimate the difficulty of the material that they recommend to library patrons. There appears to be little doubt but adults of limited reading ability would do a larger amount of reading were adequate materials available for their use.

If the attempt to furnish adults with suitable reading material is to be anything more than sheer guesswork, it is apparent that there exist two fundamental needs. One is for a scientific technique by means of which librarians and teachers of adults can determine what existing materials are of appropriate difficulty for readers of known abilities. The other is for an objective method whereby writers and publishers can determine whether particular modes of expression are such as will offer no serious obstacles to adults whose ability to read is limited. The present study is undertaken in recognition of these needs.

The purpose of the study.- This study attempts to find out (1) whether there are elements within materials themselves which may be appropriately designated significant indicators of difficulty; and (2) how these elements may be utilized in a more accurate estimate of the difficulty of specific reading materials.

Major steps of procedure.- The six major steps of the study may be summarized as follows:

1. A list of elements representing variants in the expression of adult reading material was compiled, from which the most promising were chosen as potential elements of difficulty to use for experimentation.

2. Books, magazines, and newspapers, representing adult

¹L. L. Dickerson, Libraries and Adult Education, p. 58. Report of a Study Made by the American Library Association, Chicago: American Library Association, 1926.

reading materials of a general character were examined for the purpose of obtaining desirable sources of test items to be used in the investigation.

3. From these selected materials, two reading tests were devised which were used in measuring adult comprehension of fiction and non-fiction of a general sort.

4. The tests were given to groups of adults of various levels of reading ability for the purpose of obtaining a criterion of difficulty for each test item, in terms of the average reading score.

5. Correlations were computed between a series of expressional variants contained in the test items, and the criterion of difficulty for each item, as determined by the reading tests. Evidence obtained from these correlations was used to identify elements of difficulty in adult reading materials.

6. A technique was devised in the form of a regression equation, by means of which the difficulty of general reading materials can be estimated objectively.

Scope of the investigation.- This study is limited to a determination of elements of difficulty inherent in reading materials, that is, elements of expression used by an author in presenting his ideas. Materials are restricted to those that are apparently sought by the general reader without relation to his professional and business interests. They include fiction and non-fiction of a general, informational character.

Defining Potential Elements of Difficulty

Expressional variants in adult reading materials.- The first step in the study was the compilation of a list of variants in written expression which were believed to bear some relationship to difficulty. Such a list seemed essential if the ultimate identification of elements of difficulty was to be as comprehensive as possible. The list as shown in Table I was obtained from findings of related studies, from examination of many kinds of reading material, and from suggestions made by teachers of adult classes, foreign-born adults in Americanization groups, and professors of English composition. Each suggested variant was examined for meaning, for promise of relationship to difficulty, and for objectivity of measurement. Sixty-four variants, found to be open to objective definition and measurement, were retained as potential elements of difficulty for use in experimentation. They are designated by asterisks in Table I. A brief description of these elements is given in the following paragraphs.

Easy and hard words.- In accordance with the plan followed in a related study in the field of technical reading, a word was

TABLE I

EXPRESSIONAL VARIANTS SUGGESTED AS POTENTIAL ELEMENTS
OF DIFFICULTY IN ADULT READING MATERIAL

Number	Expressional Variant
Properties of Words	
* 1	Number of easy words
* 2	Per cent of easy words
* 3	Number of different hard words
* 4	Number of words not known to 90 per cent of sixth-grade children
* 5	Per cent of words not known to 90 per cent of sixth-grade children
* 6	Number of different words
* 7	Per cent of different words
* 8	Per cent of monosyllables
* 9	Per cent of bisyllables
* 10	Per cent of polysyllables
* 11	First person pronouns
* 12	Second person pronouns
* 13	Third person pronouns
* 14	First, second, and third person pronouns
15	Nouns (proper, abstract, and concrete)
16	Verbs
17	Adjectives (descriptive, limiting)
18	Articles (definite, indefinite)
19	Adverbs (time, place, manner)
20	Conjunctions (coördinate, subordinate)
21	Interjections
22	Prepositions
* 23	Content words
* 24	Structural words
25	Image-bearing words
26	Non-image-bearing words
* 27	Words beginning with w
* 28	Words beginning with h
* 29	Words beginning with b
* 30	Words beginning with t
* 31	Words beginning with e
32	Words associated with adult living
33	Words expressing abstractions
* 34	Asides, appositives, parenthetical expressions
* 35	Dialect words
* 36	Archaic words, rare words
* 37	Local expressions and coined words
38	Poetic and highly literary words
* 39	Idiomatic expressions
* 40	Proper nouns of mythology and history
41	Technical words
Properties of Sentences	
* 42	Average sentence-length in words
* 43	Average sentence-length in syllables
* 44	Maximum syllabic sentence-length
* 45	Minimum syllabic sentence-length
* 46	Range of syllabic sentence-length
* 47	Number of simple sentences

TABLE I (Continued)

Number	Expressional Variant
* 48	Per cent of simple sentences
* 49	Number of compound sentences
* 50	Per cent of compound sentences
* 51	Number of complex sentences
* 52	Per cent of complex sentences
* 53	Number of compound-complex sentences
* 54	Per cent of compound-complex sentences
* 55	Number of compound and compound-complex sentences
* 56	Clauses introduced by subordinate conjunctions
* 57	Clauses introduced by conjunctive adverbs
* 58	Clauses introduced by relative pronouns
* 59	Prepositional phrases
* 60	Infinitive phrases
* 61	Prepositional and infinitive phrases
* 62	Explicit thought-statements
* 63	Implicit thought-statements
* 64	Static predication of thought-statements
* 65	Dynamic predication of thought-statements
* 66	Figures of speech
Properties of Paragraphs or Entire Selections	
* 67	Number of words in a selection
* 68	Number of sentences in a selection
* 69	Number of ideas in a selection
* 70	Direct discourse
* 71	Indirect discourse
72	Scenic narration
73	Dramatic narration
74	Physical associations
75	Psychic associations
* 76	A sequence of ideas to show effects
* 77	A sequence of ideas to show causes
* 78	Deductive presentation of a thought
* 79	Inductive presentation of a thought
* 80	Enumeration of ideas to show cause and effect
* 81	Repetition of an idea for emphasis
* 82	Development of an idea from a topic sentence

*Retained as a potential element of difficulty for further study.

here assumed to be "easy," regardless of its length, if it is familiar, and "hard," if unfamiliar.¹ Familiarity was established by the presence of a word in a composite list of the most frequent words in the Teachers' Word Book, by Thorndike and the Word List

¹Edgar Dale and Ralph W. Tyler, "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability," p. 18. Unpublished Investigation, Bureau of Educational Research, Ohio State University, Columbus, Ohio.

of the International Kindergarten Union.¹ Seven hundred and sixty-nine words are common to both lists. Since the occurrence of a word in the Thorndike list indicates that it is among the thousand words most frequently encountered in general reading, and since its presence in the kindergarten list is evidence of its familiarity to children before they enter first grade, the conclusion seemed warranted that the word affords no difficulty for adults. The degree of "hardness" implied in any word not classified as "easy" was determined by the element described next.

Words not known to 90 per cent of sixth-grade pupils.- The inclusion of this element as a probable index of difficulty had its origin in a study of the familiarity of 8000 words to 7878 children in the elementary school.² Since Dale found a correlation of .65 between adult scores and sixth grade scores, it seemed likely that a word not known to 90 per cent of sixth grade pupils would present difficulty to adults of limited reading ability.

Different words.- Claims have been made that range of vocabulary, or vocabulary diversity, as measured by the per cent of different words in a given selection is an indicator of difficulty.

Syllabic length of word.- Evidence found in Bear's study of the comparative difficulty of monosyllables and polysyllables in children's reading led to the conclusion that long words are generally more difficult than short words.³ It was assumed that a similar relationship might be found between length of words and reading difficulty for adult readers of limited ability.

Personal pronouns.- Other investigators have determined the relationship between the number of personal pronouns and reading difficulty, the assumption being that the informality with which material is written increases its simplicity and consequently the ease with which it is comprehended. In this study, personal pronouns in the first, second, and third person were taken separately as potential indices of difficulty, and were then combined into a single element.

Content and structural words.- The hypothesis was held that content words, namely, nouns, verbs (except copulative and auxiliary), descriptive adjectives and adverbs, tend to influence

¹Edgar Dale, "A Comparison of Two Word Lists," Educational Research Bulletin, X (December 9, 1931), 484-89.

²Edgar Dale, "Familiarity of 8000 Common Words to Pupils in the Fourth, Sixth, and Eighth Grades," Unpublished study, Bureau of Educational Research, Ohio State University, Columbus, Ohio. Pp. 84.

³Mata V. Bear, "The Length of Word as an Index of Difficulty in Silent Reading," p. 54. Unpublished Master's Thesis, Department of Education, University of Chicago, 1927.

difficulty by their intrinsic power of conveying ideas. On the other hand, structural words, namely, pronouns, copulative and auxiliary verbs, limiting adjectives, prepositions, conjunctions, and interjections, were believed to bear a negative relationship to difficulty, since their scope of meaning is more nearly constant than is generally true of content words.

Words beginning with w, h, b, i, and e.- That words whose initial letter is w, h, b, i, and e bear some relationship to difficulty was reported by Lewerenz, and later by Dale and Tyler.¹ The former found that words beginning with w, h, and b occurred with comparative frequency in easy material, while there were relatively few words beginning with i and e. In difficult reading material the situation was reversed. Dale and Tyler found significant relationship between words beginning with i and reading difficulty, but slight relationship between words beginning with other critical letters and difficulty of technical selections.

Asides, appositives, and parenthetical expressions.- Since extremely elementary materials tend to utilize words, phrases, and clauses of co-ordinate rank, it was believed that subordination of certain elements in a thought-unit would present difficulties for the reader of immature habits. Such elements, whose subordination is indicated by commas, parentheses, or dashes, were accordingly considered.

Other variants pertaining to words.- Other variants that were believed to influence difficulty were words whose meanings have acquired environmental familiarity for some persons, but have remained unknown to others. These were dialect words, archaic expressions and rare words, local expressions and coined words, and proper nouns of history and mythology.

Length and kind of sentence.- Although some investigators have found that the effect upon the reader is the same, whether a sentence is long and involved, or short, simple, and emphatic, the weight of evidence points toward a positive relationship between length and complexity of sentence and reading difficulty. Length of sentence was here measured by average number of words, average number of syllables, the syllabic extremes of all sentences in a selection, and the syllabic range between the longest and the shortest sentence. Complexity of sentence-structure was measured by the occurrence of simple, compound, complex, and compound-complex sentences. Complex sentences were further analyzed according to the type of subordination involved.

¹Alfred S. Lewerenz, "Measurement of the Difficulty of Reading Materials," Los Angeles Educational Research Bulletin, VIII (March, 1929), 11-16; Dale and Tyler, op. cit., p. 17.

Kinds of phrases.- In the small and concrete vocabulary of very simple people few phrases have been observed. This observation led to the assumption that phrases complicate context, and are, therefore, elements of difficulty. Prepositional and infinitive phrases were considered individually and then jointly.

The influence of predication.- The freedom and ease of modern writing, accomplished in part by a disregard for predication, provoked the question as to whether an explicit sentence, in which thought relationships are fully expressed, influences difficulty of comprehension in the same degree as does the implicit sentence, whose thought relationships are only suggested.

Consideration of explicit sentences led to a second question as to whether dynamic predication, which implies a change of condition, and static predication, which merely associates ideas without implying change, are related to difficulty in the same or in a different direction. Since a study of the evolution of human speech shows that in primitive speech the dominant mode of predication employs action words, it was believed that dynamic predication might promote understanding.

Figures of speech.- On the assumption that figurative language is related to difficulty of comprehension, five figures of speech most common in general material, were studied. They were: simile, metaphor, metonymy, synecdoche, and personification.

Length of a selection.- There is evidence that brevity of statement is essential to comprehension for the inexperienced reader and that it is the better reader who chooses large books, long chapters, or detailed treatises. Length of a selection was accordingly included among potential elements of difficulty. It was measured in three ways, by words, sentences, and ideas.

Direct and indirect discourse.- The use of direct and indirect discourse was believed to increase ease of comprehension. That the first is more closely related to ease than the second seemed likely, on the assumption that as the reader temporarily projects himself into another personality and expresses his sentiments and ideas, he acquires a clearer understanding of their meanings.

Techniques of paragraph development.- The last seven variants in Table I relate to the author's technique of developing paragraphs by a sequence of ideas to show effect, a sequence of ideas to show cause, deductive presentation of an idea, inductive presentation of an idea, enumeration of ideas to show both cause and effect, repetition of an idea for emphasis, and development of an idea from a topic sentence. It was believed that whereas some of these techniques promote comprehension, others retard it.

Uses of compiled elements.- Some of the elements listed in

Table I were used in the selection of materials as sources of test items. Others were used in ranking test items according to probable levels of difficulty. Finally, all of them were used in a quantitative study of test items for the purpose of identifying elements which are related to difficulty.

Selecting Reading Materials

Two conditions were held essential to the selection of serviceable and reliable materials to use in determining elements of difficulty. The first condition pertained to the representativeness of the materials. If tests derived from them were to measure the comprehension of adult readers, it appeared desirable that the materials should represent not only what is written for adults, but what is actually read by them.

The second condition related to the range of difficulty presented by the materials. It is entirely possible to choose materials so difficult that test items derived from them will not measure the comprehension of adults whose reading ability is limited. On the other hand, materials may be simple that an inadequate measure of comprehension will be obtained for the more able readers. An attempt was made, therefore, to choose representative materials with a scope of difficulty so wide that tests derived from them may be adapted to readers of different abilities.

Selecting representative materials.- Since detailed studies of the reading activities of adults reveal the fact that 50 per cent of them read books, 75 per cent read magazines, and 95 per cent or more read newspapers, these three types of material may be thought to represent in a general way what adults are reading.¹

In Table II are listed the books chosen for study. They include books of fiction and non-fiction which were most popular from January, 1929, until June, 1931, as determined by monthly book lists in The Bookman and booksellers' reports in The Publishers' Weekly; "best" standard books;² standard fiction in original and adapted versions; a literary experiment in Basic English; a series of reading text books for middle grades; and the Bible.

¹William S. Gray and Ruth Monroe, The Reading Interests and Habits of Adults, p. 262. New York: Macmillan Company, 1929.

²H. W. L., "Million Books and Best Books," The Golden Book, VIII (August, 1926), 382-83.

TABLE II

NAMES OF TWENTY-NINE BOOKS SELECTED AS SOURCES OF TEST ITEMS,
WITH CLASSIFICATION AND AUTHOR OF EACH

Classification	Book	Author
Popular Novels	All Quiet on the Western Front	Erich Maria Remarque
	Angel Pavement	J. B. Priestley
	Cimarron	Edna Ferber
	Peder Victorious	O. E. Rolvaag
	Roper's Row	Warwick Deeping
	Scarlet Sister Mary	Julia Peterkin
Popular General Books	White Oaks of Jalna	Mazo de la Roche
	Art of Thinking, The	Ernest Dimnet
	Byron	Andre Maurois
	Elizabeth and Essex	Lytton Strachey
	Henry the Eighth	Francis Hackett
	A Preface to Morals	Walter Lippman
"Best" Standard Books	Story of San Michele, The	Axel Munthe
	Tragic Era, The	Claude G. Bowers
	Adventures of Huckleberry Finn, The	Mark Twain
	Last of the Mohicans, The	James Fenimore Cooper
	Scarlet Letter, The	Nathaniel Hawthorne
	Tales	Edgar Allen Poe
Classics, Original Form	Moby Dick	Herman Melville
	Robinson Crusoe	Daniel Defoe
Classics, Adapted Form	Silas Marner	George Eliot
	Moby Dick	Sylvia Chatfield Bates
Adaptation in Basic English	Robinson Crusoe	Ettie Lee
	Silas Marner	Ettie Lee
School Readers for Middle Grades	Carl and Anna	Frank Leonhard
	Reading and Living, Book I	Hill, Lyman & Moore
	Reading and Living, Book II	Hill, Lyman & Moore
	Reading and Living, Book III	Hill, Lyman & Moore
	Bible	King James Version

Magazines chosen as sources of test items appear in Table III. They were restricted to general magazines whose circulation figures equal or exceed 125,000, as sworn to by the Audit Bureau of Circulation.¹ The selection of a larger number of magazines than of other classes of material is explained by the fact that they had been used by the writer in an earlier study, and considerable data relative to their content were already compiled.² Dates

¹N. W. Ayer & Sons, Directory of Newspapers and Periodicals for 1931. Philadelphia: N. W. Ayer & Sons, 1931. Pp. 1386.

²Bernice E. Leary, "An Analytical Study of Certain Elements

of issue for magazines used in the investigation ranged from December, 1930, to April, 1931.

TABLE III
MAGAZINES SELECTED AS SOURCES OF TEST ITEMS

Name of Magazine	
Ace-High Magazine Aces (Fiction House Group) All Star Detective Stories All-Story American Home American Magazine, The Atlantic Monthly, The Better Homes and Gardens Blue Book Magazine, The Capper's Farmer Clues College Humor Collier's Magazine Country Gentleman Delineator Detective Story Magazine Dream World Farmer's Wife Field and Stream Gentlewoman, The Golden Book Magazine, The Good Housekeeping Hearst's International Cosmopolitan Magazine Holiday Holland's Magazine House and Garden Household Magazine, The Judge Ladies Home Journal, The Liberty Literary Digest, The McCall's Magazine Mother's Home Life Motion Picture Classic	Motion Picture Magazine, The Nation's Business National Farm Journal, The National Geographic Magazine, The Needlecraft Magazine New Movie Magazine, The Outdoor Life Pathfinder, The People's Popular Monthly Photoplay Magazine Pictorial Review Picture Play Magazine Popular Mechanics Magazine Popular Science Monthly Radio News Ranch Romances Red Book Magazine Review of Reviews Rotarian, The Saturday Evening Post Science and Invention Screen Book Magazine Screenland Short Stories Sky Riders Successful Farming Time True Detective Mysteries True Romance True Story Magazine Vogue Woman's Home Companion Woman's World World's Work

Table IV lists fifteen newspapers selected as sources of
in the Reading Material of Adult Magazines," Unpublished Master's
Thesis, Dep't of Education, University of Chicago, 1931. Pp. 122.

test items. In general, daily newspapers in five large cities were chosen if they had the widest circulation among either the morning or evening editions. For Chicago, both a morning and an evening paper were chosen. Village and rural publications were chosen at random from towns in the Middle West whose population figures range from 6,140 to 402. All of the editions used were dated between June 27, 1931, and August 13, 1931.

TABLE IV
NAMES OF FIFTEEN NEWSPAPERS SELECTED AS SOURCES OF TEST ITEMS

Classification	Newspaper
Metropolitan	Boston Post
	Chicago American
	Chicago Daily Illustrated Times (Picture Newspaper)
	Chicago Daily Tribune
	Christian Science Monitor
	New York Evening Graphic
	New York Journal
	New York News (Picture Newspaper)
	Los Angeles Times
	Philadelphia Evening Bulletin
Village	Edmore Herald News (North Dakota)
	Jefferson Banner (Wisconsin)
	Olney Daily Mail (Illinois)
	Riceville Recorder (Iowa)
Rural Digest	Capper's Weekly (Kansas)

Ranking materials for probable difficulty.- In order that test items used in the study would provide a valid measure of comprehension for adults of different abilities, materials from which they were developed were roughly classified as "easy," "average," or "difficult" by the following plan:

1. Each book, magazine, and newspaper was analyzed for the occurrence of four "promising" elements: syllabic length of word, vocabulary diversity, or per cent of different words, syllabic length of sentence, and type of sentence. A detailed description of the technique involved in this procedure is given in the writer's earlier study.

2. Data obtained from the analysis of each element in the different materials were then summarized statistically, to show the range, the central tendency, and the variability of occurrence of each element in the three classes of material.

3. Books, magazines, and newspapers were roughly classified as "easy," "average," or "difficult," as determined by the deviation of the mean occurrence of each element in their content from the mean occurrence for all materials of their class. "Easy" was defined as an extreme variation from the central tendency for a particular element in the direction of simplicity; "difficult," as extreme deviation in the opposite direction, that is, toward complexity; and "average," as coincidence or approximate coincidence with the central tendency.

4. A final classification was made as shown in Table V, representing the composite classifications by all four elements.

TABLE V
CLASSIFICATION OF READING MATERIAL INTO GROUPS
OF PROBABLE DIFFICULTY

	Easy	Average	Difficult
Book	All Quiet on the Western Front Carl and Anna Huckleberry Finn Moby Dick (Adapted) Robinson Crusoe (Adapted) Roper's Row Silas Marner (Adapted)	Angel Pavement Bible Byron Cimarron Elizabeth and Essex Henry the Eighth Moby Dick The Scarlet Letter	Art of Thinking Last of the Mohicans A Preface to Morals Tragic Era, The
Magazine	Aces All-Story Dream World	American Better Homes and Gardens Collier's Cosmopolitan Liberty Saturday Evening Post Screenland	Atlantic Monthly Radio News Review of Reviews
Newspaper	Capper's Weekly	Boston Post Jefferson Banner Riceville Recorder	Christian Science Monitor Los Angeles Times

Since a similar classification as "easy," "average," or "difficult" by more than two individual analyses was requisite to a

final ranking into any particular group, many materials could not be classified in Table V. Although materials shown in the table were considered the most desirable sources of test items, the omission of a book, magazine, or newspaper did not preclude its usefulness for that purpose, since usefulness was conceded on the grounds of its representativeness and its general direction of difficulty.

Constructing Adult Reading Tests

Measurement of comprehension.- That comprehension is not a single unitary process is a generally accepted theory. It is, rather, a blending of many processes, whose totality represents understanding. Because comprehension has not been satisfactorily resolved into its component processes, now one process and now another is taken as a measure of comprehension, as test-makers face the baffling questions of what to measure and how to measure.

Two aspects of comprehension directed the construction of the tests used in this study: (1) the ability to get the "sense" of what is read in the form of a general impression, evidenced by the statement of a generalization or summary, and (2) the ability to recall specific elements in material, evidenced by a knowledge of the facts presented therein. These aspects were held of first importance among the major outcomes inferred from reading general material for a variety of purposes: "for fun," "for emotional satisfaction," "to acquire general information," "to gather specific information," "to extend experience," "to improve one's literary appreciation," "to satisfy one's curiosity," "to forget."¹ Since the ability to grasp the essential meaning seems to depend somewhat upon the ability to recall specific elements contained in the selection, it was assumed that any technique devised to measure the first outcome would indirectly measure the second. Other major outcomes were not tested except insofar as they may be correlated with the abilities measured.

The multiple-response test adapted for this study was composed of a number of disconnected paragraphs from reading materials selected as sources of test items. In order to increase the reliability of the measure of comprehension two reactions were required. The first was the identification of the best summary of the

¹Gray and Monroe, op. cit.; The Twenty-fourth Year Book; William S. Gray, "The Importance of Intelligent Silent Reading," The Elementary School Journal, XXIV (January, 1924), 349-52; Rhey B. Parson, "A Study of Adult Reading," Unpublished Master's Thesis, Dept. of Education, University of Chicago, 1923. Pp. 124; Wilda Lee Montgomery, "The Investigation of the Uses of Recreatory Reading," University of Pittsburgh School of Education Journal,

paragraph in a series of five statements relating to the paragraph. The second was the recognition in the series of a statement not in the paragraph.¹ Reactions were indicated in the manner prescribed by the following passage, which is quoted from the set of directions accompanying the tests:

"You are to read each paragraph carefully. Then read the the five sentences below the paragraph. Put a CHECK MARK (✓) before the sentence that tells best what the paragraph said. One of the five sentences tells something that is wrong or that is not in the paragraph. Mark that sentence with a ZERO (0). Do not mark the other three sentences."

The nature of the test is apparent from the following sample paragraph.

"Sample Paragraph

"For nine years I have watched, during the spring and summer months, a nighthawk that lives in the daytime on a big limb near the end of my garden. He can usually be found at the same spot from April to September. It may be that a member of a second or third generation now lives there; but it is truly a family tree. I often wonder why a bird becomes attached to one place, unless it is because it is contented there. Explain it as you will, contentment results from peace of heart, and the home-loving heart is usually at peace.

"Paragraph Summary

- "1.--Ever since the nighthawk built his nest in my garden he has lived there all the year round.
- "2.--The tree where the nighthawk lives is truly a family tree.
- "3.--The nighthawk can usually be found on the same limb during spring and summer.
- "4.--The nighthawk must be contented in my garden for he has lived there from April to September for nine years.
- "5.--The same nighthawk or perhaps a member of his family has lived in my garden during spring and summer for nine years."

Twenty-four paragraphs of fiction are contained in Form 1, and the same number of paragraphs of non-fiction, in Form 2.

Selecting test items.—Test items were chosen from a large number of sample paragraphs selected from representative books, magazines, and newspapers previously classified as "easy," "average," or "difficult." Criteria of selection were determined by the nature of the test, the outcomes to be measured, the kind of adults to be tested, and the specific elements whose difficulty

IV (March-April, 1929), 90-91.

¹The technique was suggested by Dale and Tyler, op. cit.

was to be established. The criteria were stated in the form of the following principles:

1. Each item must be independent of the content preceding and following it.
2. Each item must manifest a completeness, indicated by a simple unit of thought.
3. Each item must contain as many variants in expression as possible.
4. Each item must be brief, approximately 100 words in length.

Over forty paragraphs of fiction and of non-fiction were chosen from books, magazines, and newspapers. These paragraphs were then ranked in order of probable difficulty as determined by the occurrence of ten elements: per cent of easy words, number of different hard words, number of words not known to 90 per cent of sixth-grade pupils, per cent of monosyllables, per cent of polysyllables, number of prepositional and infinitive phrases, number of first, second, and third person pronouns, average sentence-length in syllables, and number of e words. From these rankings paragraphs which held consistently to a particular level of relative difficulty among all paragraphs were selected as test items. Levels of difficulty were here defined in terms of deviation in the number or per cent of occurrence of an element in the direction of simplicity or complexity. Items of fiction and of non-fiction were then ranked from 1 to 24 in progressive stages of probable difficulty for two test forms.

Preparing responses for test items.- As stated previously, five responses were formulated for each item. Three competent judges were asked to formulate independently a best and a worst summary statement for each item, to evaluate the statements in conference, and to formulate three others in approximately normal distribution between the extremes. Each summary was to be stated in language less difficult than that of the test item in order to prevent later assignment of a degree of difficulty to an item that was truly the product of the difficulty of its responses. Two groups of students at the University of Chicago evaluated the responses for the purpose of determining how well the responses had been scaled from best to poorest, and the extent to which the wording of the responses involved difficulty. Their reactions were tabulated and used as a basis for revision of the test responses.

Evaluating test responses.- The next step was to determine the value of each response, that is, the amount of credit that an individual should receive for marking it the best or the poorest. There was believed to be greater fairness attached to giving credit for every response than giving credit for only the summaries established as best and poorest. Since an attempt had been made to

scale the responses according to the goodness of their stated summaries, it was assumed that for every item there was a second response closer to the best than any other; and also a second response closer to the poorest than any other. There was presumably an intermediate response as close to the best as to the poorest. If an individual should choose a second-rate response, he should receive credit indicating better comprehension than would have been indicated by choosing the intermediate response as "best" or "poorest." Although assigning credit for every response, which would account for varying degrees of merit, was less expeditious than assigning credit for only the best and the poorest summaries, it seemed a more valid measure of comprehension.

Nine persons who had had considerable experience with tests were asked to rate the five responses for each test item. They were directed to signify the best response by a rating of 1, the poorest response by a rating of 5, and other responses according to their judgment. They were also directed to indicate for each item the degree of difficulty experienced in selecting the best and the poorest responses by recording E for easy, M for medium, and H for hard. From the independent ratings of each response by nine judges, a mean value was calculated which represented the credit received by any individual checking that response as best or poorest. When the reliability of the mean values was calculated in terms of $\sigma'_M = \frac{\sigma(\text{dis})}{N}$, it was found that the closest agreement between mean values and true values occurred in the case of best and poorest summaries. The fact that all standard errors were relatively low, however, indicated that all values might be considered generally reliable.

The reliability of the composite judges for both forms of the test ranged from .994 to .970, when responses with lowest and highest variability in judgment represented the range in reliability.¹ For purposes of this study, nine judges were considered

¹The formula for solving the average intercorrelations made use of actual ratings without computing any of the individual correlation coefficients:

$$\bar{r} = \frac{1}{Nn(n-1)} \left\{ \sum \left[S \left(\frac{X}{\sigma} \right) \right]^2 - N \left[S \left(\frac{M}{\sigma} \right) \right]^2 \right\} - \frac{1}{n-1}, \text{ in which } N = \text{the}$$

number of responses; n = the number of judges; S means a summation of each response of $\frac{X_i}{\sigma}$, where $i = A, B, C, \dots, I$ judges; $\frac{X}{\sigma} = \frac{\text{rating}}{\sigma}$

for one judge. See: Harold A. Edgerton and Herbert A. Toops, "A Formula for Finding the Average Intercorrelation Coefficient of Unranked Raw Scores Without Solving Any of the Individual Inter-

reliable, since perfect reliability of composite ratings is possible only by using an infinite number of judges.

Rating the difficulty of responses.- A somewhat arbitrary method, designated the 2-1-0 method, was used to give numerical values to the difficulty ratings of the responses, indicated E, M, and H by the judges. Each H rating was given 2 points; each M rating was given 1 point; and each E rating, 0 points. All the ratings of the nine judges were tabulated by points for the best and the poorest response in each paragraph. It was then possible to determine whether variations in difficulty of choosing the best and the poorest responses were of enough significance to require weighting the value of the selections. A deviation of four points in either direction from the mode marked a limit above or below which the difficulty of a response was considered distinct from the difficulty of other responses in the test. Approximately all of the best and poorest responses in both tests were found to deviate relatively little from the mode. In no item was weighting of responses considered necessary. It was apparent, therefore, that difficulty due to wording of the responses was generally insignificant.

Identifying Elements That Influence Difficulty

Obtaining a criterion of difficulty.- Elements of difficulty were identified in general reading material by the method of correlation, which established the relationship between different elements occurring in each test item and the criterion of difficulty for that item. As was pointed out earlier, the criterion of difficulty used in the present study was the average reading score made on the test items by organized groups of adults. A criterion of difficulty was obtained for three classes of readers. One criterion represented the average reading score on each test item of 756 adults of different levels of reading ability. A second represented the average reading score of the "best" readers, defined here as those individuals whose combined scores on Form 1 and Form 2 of the reading tests ranked above the third quartile. A third criterion represented the average reading score of the "poorest" readers, that is, of those individuals whose combined scores

correlations," Journal of Educational Psychology, XIX (February, 1928), 131-38. Formula for prediction is

$$r_{nn} = \frac{r_{1I}}{1 + (n-1)r_{1I}}. \text{ See: Karl J.}$$

Holzinger, Statistical Methods for Students in Education, p. 169. Boston: Ginn & Co., 1928.

ranked below the first quartile. By the use of three criteria, it was possible to identify elements of difficulty for general readers, for "best" readers, and for "poorest" readers.

The groups tested.- The tests were given to groups of adults who were attending school and who were organized at different educational levels, and to non-school groups of a relatively heterogenous composition. It was hoped that the former groups would give an adequate representation of adults of limited reading ability, and that the latter groups would be sufficiently diversified as to abilities, interests, and educational background to give a fair representation of the general reading public.

Twenty-nine groups were used in the entire study. Of these, twenty-one groups comprising 756 subjects were used in the major part of the study to obtain a criterion of reading difficulty. Use was made of the remaining eight groups in establishing the reliability of the findings of the study. Table VI presents data pertaining to the general characteristics of the twenty-one groups who were tested first.

Additional facts pertaining to the individuals tested may be summarized as follows:

1. The age of the 756 subjects ranged from 15 years to 50 years, the mean age being 22.7 years.
2. Educational training, as measured by last grade attended in school, ranged from third grade, reported by one individual, to beyond twelfth grade, reported by 246. The mean was calculated at grade 11.01, by interpreting "beyond 12" as 14.
3. Of the 750 subjects who answered the question pertaining to attendance at night school, 660 reported from no attendance to three months, while three claimed a period of attendance between five and six years in length. The mean attendance was 3.1 months.
4. Ability in reading as measured by the Monroe test of comprehension ranged from grade 16.95 to grade 1.95, with a mean of 9.68.
5. Amount of time devoted to reading per day represented a crude estimate. Two subjects reported that they did no reading, whereas seventy-three recorded an average per day of more than three hours. The mean was 1.56 hours.

The testing program.- The testing program was carried out during January and February, 1932. Tests were administered by the writer to groups living in Illinois, and by a supervisor or psychologist to other groups. Two periods, approximately an hour in length, were usually needed. The Adult Reading Tests were administered in accordance with relatively simple directions.

TABLE VI

NAME, RACE, CLASSIFICATION, GRADE PLACEMENT, GEOGRAPHICAL LOCATION, AND SIZE OF TWENTY-ONE GROUPS
OF ADULTS TESTED BY BOTH FORM 1 AND FORM 2 OF THE ADULT READING TEST

Number	Name of Group	Race	Classifi- cation	Grade Place- ment	Geographical Location	Number of Cases
1	Dante Adult Day School	Mixed	Elementary	8	Chicago, Ill.	12
2	Engelwood Night School	Mixed	Elementary	7-8	Chicago, Ill.	30
3	J. Sterling Morton Night School.	Mixed	Elementary	7-8	Chicago, Ill.	25
4	Glenn Street Evening School. . . .	White	Elementary	6	Anderson, S. C.	13
5	Berea Foundation	White	Junior High	7-9	Berea, Ky.	90
6	A. & M. Model School	Black	Junior High	7	Tallahassee, Fla.	11
7	A. & M. Model School	Black	Junior High	8	Tallahassee, Fla.	14
8	A. & M. Model School	Black	Junior High	9	Tallahassee, Fla.	23
9	A. & M. Model School	Black	Senior High	10	Tallahassee, Fla.	28
10	A. & M. Model School	Black	Senior High	11	Tallahassee, Fla.	17
11	A. & M. Model School	Black	Senior High	12	Tallahassee, Fla.	28
12	Berea Academy.	White	Senior High	9-12	Berea, Ky.	97
13	Martha Berry	White	Senior High	9	Rome, Ga.	44
14	Martha Berry	White	Senior High	11-12	Rome, Ga.	14
15	Rabun-Gap Nacoochee Mt. School .	White	Senior High	10	Rabun Gap, Ga.	23
16	Rabun-Gap Nacoochee Mt. School .	White	Senior High	11	Rabun Gap, Ga.	17
17	Berea College.	White	College	Fr.-Sr.	Berea, Ky.	98
18	Martha Berry	White	College	Fresh.	Rome, Ga.	60
19	Martha Berry	White	College	Juniors	Rome, Ga.	29
20	Hinsdale Parent-Teachers' Ass'n.	White	Non-School		Hinsdale, Ill.	38
21	Congress Park Parent-Teachers' Association	White	Non-School		Congress Park, Ill.	45
Total						756

Every effort was made to establish harmonious relations between the test administrator and the subjects, in order that the results should be as reliable as possible.

Scoring test items.- For any response designated as poorest by "0," a score was given which represented the mean rating of nine judges. From this score was subtracted the value of the response designated as best by " $\sqrt{}$." Hence the score on any item was the difference between the value assigned to the poorest response and the value assigned to the best.¹ Had all nine judges been in perfect agreement as to the rating of the best and the poorest responses, the best response would have always a value of 1, and the poorest, a value of 5. A reader who checked both responses correctly would receive a score representing the difference between 5 and 1, or 4. A reader who checked either or both responses incorrectly would receive a score less than 4. Actually, however, an average value was frequently not an integer, but a decimal, 4.9, 1.2, and so on.

To illustrate how the method of scoring was used, a test item from Form 2 will be taken as an example. In this item, the best summary was given a value of 1.1; the second best, 1.6; the third best, 3; the fourth best, 3.6; and the poorest, 5. If a reader checked both the best and the poorest responses correctly, his score would be 5 minus 1.1, or 3.9, the highest possible score on that item. If he checked the worst response correctly, but checked the second best as best, his score would be 5 minus 1.6, or 3.4. If he checked the worst response correctly, but chose the third best as best, his score would be 5 minus 3, or 2. If he made the worst possible choice, that is, if he checked the poorest response as best and the best as poorest, his score would be 1.1 minus 5, or -3.9. The possible range of scores for any test item, therefore, is from a theoretical + 4 to -4, with many possible intervening scores expressed either integrally or decimally.

Scores on each test item for all individuals were entered on data sheets, and an average score computed for each item. This average score represented the criterion of difficulty on a given item for 756 subjects. "Best" and "poorest" readers were then identified, and their average reading scores computed for each test item. Average scores for the three classes of readers are shown in Tables VII and VIII. It should be remembered, in reading these data, that a high average score indicates an easy selection, and a low average score, a difficult selection.

¹The writer is indebted to Professor Ralph W. Tyler of the Bureau of Educational Research, Ohio State University, Columbus, Ohio, for this method of scoring.

TABLE VII

CRITERIA OF DIFFICULTY FOR ITEMS ON FORM 1, REPRESENTED BY THE
AVERAGE SCORE OF ALL READERS, "BEST" READERS,
AND "POOREST" READERS

Test Item	Average Reading Score			Test Item	Average Reading Score		
	All Readers	"Best" Readers	"Poorest" Readers		All Readers	"Best" Readers	"Poorest" Readers
1	2.6	3.7	1.1	13	2.4	3.4	1.1
2	2.3	3.0	1.5	14	2.3	3.1	1.0
3	2.1	3.3	0.5	15	2.4	3.4	1.2
4	2.6	3.2	1.5	16	1.9	2.8	0.7
5	2.5	3.2	1.5	17	2.5	3.6	1.0
6	3.0	3.7	2.0	18	1.9	3.4	0.1
7	2.6	3.8	1.3	19	2.4	3.2	0.9
8	2.2	3.2	1.0	20	2.1	3.3	0.7
9	2.0	3.0	0.6	21	1.9	3.2	0.6
10	2.4	3.8	0.4	22	2.0	3.5	0.5
11	2.1	3.3	0.7	23	1.9	3.1	0.7
12	1.9	3.4	0.2	24	1.7	3.2	0.4
Range					3.0-1.7	3.8-2.8	2.0-0.1
Mean					2.2 $\pm$.070	3.3 $\pm$.066	0.9 $\pm$.058
S.D.					.510 $\pm$.050	.480 $\pm$.047	.424 $\pm$.041

TABLE VIII

CRITERIA OF DIFFICULTY FOR ITEMS ON FORM 2, REPRESENTED BY THE
AVERAGE SCORE OF ALL READERS, "BEST" READERS,
AND "POOREST" READERS

Test Item	Average Reading Score			Test Item	Average Reading Score		
	All Readers	"Best" Readers	"Poorest" Readers		All Readers	"Best" Readers	"Poorest" Readers
1	2.9	3.5	2.1	5	2.8	3.5	1.5
2	2.6	3.5	1.2	6	2.6	3.2	1.7
3	2.4	3.2	1.7	7	1.3	2.5	0.5
4	2.6	3.2	1.6	8	0.8	1.2	0.1

TABLE VIII (Continued)

Test Item	Average Reading Score			Test Item	Average Reading Score		
	All Readers	"Best" Readers	"Poorest" Readers		All Readers	"Best" Readers	"Poorest" Readers
9	2.6	3.5	1.1	17	1.9	3.0	0.5
10	3.0	3.6	2.0	18	2.2	3.7	0.2
11	1.2	2.8	-0.2	19	1.5	3.1	0.3
12	2.5	3.2	1.3	20	1.3	3.1	0.1
13	2.3	3.3	1.0	21	0.9	2.7	-0.1
14	1.5	2.3	0.9	22	0.7	1.6	0.2
15	2.6	3.3	1.5	23	1.8	3.2	0.3
16	1.8	3.0	0.5	24	1.8	3.3	0.7
Range					2.9-0.7	3.7-1.2	2.1-0.2
Mean					2.0 $\pm$	3.0 $\pm$	0.9 $\pm$
					.097	.095	.087
S.D.					.707 $\pm$	.693 $\pm$	.632 $\pm$
					.069	.067	.062

Identifying elements of difficulty.- Obtaining a criterion of difficulty for each test item represented the first step in identifying elements of difficulty in the items. The second step was making a quantitative study of the elements in each item that might contribute to difficulty. Since only those elements were retained in Table I that are objectively measurable, their occurrence was determined by simply counting each element and expressing the result as a numerical total or in terms of per cent. Of the sixty-four elements retained for experimentation, forty-four occurred in enough items to warrant correlation; that is, in approximately half of the items.

The final step in identifying elements of difficulty was finding the relationship between the criterion of difficulty for each item and the elements in the items. Table IX presents the coefficients of correlation between the forty-four elements and the average reading score for all readers, for "best" readers, and for "poorest" readers. Although all of the correlations may seem low, their significance is to be interpreted from their relative size, rather than their absolute size. Of the correlations shown in the first column, only the first twenty-one are statistically significant, since a coefficient must be at least .27 to be three times its probable error. Actually, however, if the occurrence of an element bears even a small numerical relationship to difficulty,

TABLE IX

CORRELATIONS BETWEEN FORTY-FOUR POTENTIAL ELEMENTS OF DIFFICULTY AND THE AVERAGE SCORES OF EACH TEST ITEM FOR THREE CLASSIFICATIONS OF ADULT SUBJECTS

Number	Element	Correlation Coefficients					
		All Readers		"Best" Readers		"Poorest" Readers	
		r	P.E.	r	P.E.	r	P.E.
1	Average sentence length in words	-.522 ±	.0708	-.431 ±	.0793	-.449 ±	.0777
2	Per cent of easy words	.520 ±	.0710	.271 ±	.0902	.529 ±	.0701
3	Number of words not known to 90 per cent of sixth-grade pupils	-.513 ±	.0717	-.218 ±	.0927	-.550 ±	.0679
4	Number of easy words	.511 ±	.0719	.292 ±	.0891	.518 ±	.0712
5	Number of different hard words	-.496 ±	.0734	-.191 ±	.0938	-.543 ±	.0686
6	Minimum syllabic sentence length	-.491 ±	.0739	-.481 ±	.0748	-.406 ±	.0813
7	Number of explicit sentences	.482 ±	.0747	.380 ±	.0833	.419 ±	.0803
8	Number of first, second, and third person pronouns	.476 ±	.0753	.347 ±	.0856	.450 ±	.0776
9	Maximum syllabic sentence length	-.474 ±	.0755	-.302 ±	.0885	-.407 ±	.0812
10	Average sentence length in syllables	-.472 ±	.0757	-.289 ±	.0892	-.457 ±	.0770
11	Per cent of monosyllables	.433 ±	.0791	.166 ±	.0947	.488 ±	.0742
12	Number of sentences per paragraph	.431 ±	.0793	.337 ±	.0863	.371 ±	.0840
13	Per cent of different words not known to 90 per cent of sixth-grade pupils	-.402 ±	.0816	-.227 ±	.0923	-.409 ±	.0811
14	Number of simple sentences	.390 ±	.0825	.389 ±	.0826	.298 ±	.0887
15	Per cent of different words	-.380 ±	.0833	-.181 ±	.0942	-.441 ±	.0784
16	Per cent of polysyllables	-.380 ±	.0833	-.077 ±	.0968	-.428 ±	.0795
17	Number of prepositional phrases	-.345 ±	.0858	-.062 ±	.0970	-.370 ±	.0840
18	Number of third person pronouns	.329 ±	.0868	.150 ±	.0952	.307 ±	.0832
19	Range of syllabic sentence length	-.321 ±	.0873	-.114 ±	.0961	-.279 ±	.0898
20	Number of different words	-.291 ±	.0891	-.075 ±	.0968	-.367 ±	.0842
21	Number of infinitive and prepositional phrases	-.293 ±	.0896	-.050 ±	.0971	-.271 ±	.0902
22	Per cent of content words	-.261 ±	.0907	-.173 ±	.0944	-.232 ±	.0921
23	Per cent of structural words	.261 ±	.0907	.173 ±	.0944	.232 ±	.0921

TABLE IX (Continued)

Num- ber	Element	Correlation Coefficients					
		All Readers		"Best" Readers		"Poorest" Readers	
		r	P.E.	r	P.E.	r	P.E.
24	Number of i words	-.235 +	.0920	-.120 +	.0960	-.285 +	.0894
25	Number of Figures of speech	-.233 +	.0921	.088 +	.0966	-.377 +	.0835
26	Per cent of complex sentences	-.229 +	.0922	.359 +	.0848	-.093 +	.0965
27	Number of compound-complex sentences	.205 +	.0933	.249 +	.0913	.138 +	.0955
28	Number of infinitive phrases	.199 +	.0935	.038 +	.0972	.300 +	.0886
29	Number of first person pronouns	.194 +	.0937	.272 +	.0902	.111 +	.0962
30	Number of complex sentences	.193 +	.0937	.006 +	.0974	.255 +	.0910
31	Per cent of bisyllables	-.193 +	.0937	-.189 +	.0939	-.187 +	.0939
32	Number of h words	.183 +	.0941	.051 +	.0971	.168 +	.0946
33	Per cent of simple sentences	.180 +	.0942	.278 +	.0898	.096 +	.0965
34	Number of clauses introduced by subordinate conjunctions	-.160 +	.0949	-.198 +	.0935	-.177 +	.0943
35	Number of w words	.146 +	.0953	.045 +	.0972	.127 +	.0958
36	Number of compound and compound-complex sentences	.141 +	.0954	.194 +	.0937	.066 +	.0969
37	Number of clauses introduced by relative pronouns	.117 +	.0960	.028 +	.0973	.234 +	.0920
38	Number of asides	-.097 +	.0964	.079 +	.0967	-.211 +	.0930
39	Total number of words per paragraph	.091 +	.0965	.168 +	.0946	.060 +	.0970
40	Number of b words	-.069 +	.0969	.004 +	.0974	-.093 +	.0965
41	Number of e words	-.065 +	.0969	.198 +	.0935	-.153 +	.0951
42	Per cent of compound-complex sentences	.060 +	.0970	.114 +	.0961	.023 +	.0973
43	Number of clauses introduced by conjunctive adverbs	.054 +	.0971	-.018 +	.0973	-.019 +	.0973
44	Per cent of compound and compound-complex sentences	-.009 +	.0973	.070 +	.0969	-.075 +	.0968

nothing can seemingly justify indifference toward it, since it is probable that the relationship did not "just happen."

For all adults, twenty-three elements of difficulty were discovered whose relationship to difficulty ranged from $-.522$ for average sentence-length in words to $-.009$ for per cent of compound-complex sentences. Nineteen elements, found to correlate positively with difficulty for "best" readers ranged from $-.481$ for minimum syllabic sentence-length, to $-.018$ for clauses introduced by conjunctive adverbs. Twenty-four elements of difficulty were found for "poorest" readers, with coefficients ranging from $-.550$ for number of words not known to 90 per cent of sixth grade pupils, to $-.019$ for clauses introduced by conjunctive adverbs. It is apparent that as a single measure of difficulty, average sentence-length in words was found to be best for all readers; minimum syllabic sentence length, for "best" readers; and number of words not known to 90 per cent of sixth grade pupils, for "poorest" readers. Other facts presented in Table IX indicate that elements pertaining to sentence structure correlated more closely with difficulty for "best" readers than did other elements. In the case of "poorest" readers elements in vocabulary correlated most closely with difficulty.

A comparison of the three series of correlations shows that although an element of difficulty tends to be general, the extent of its influence varies with different classes of readers, in the direction of increasing difficulty for "poorest" and decreasing difficulty for "best" readers. One may infer that reading ability may ultimately reach a level at which comprehension is considerably less influenced by the author's mode of expression than at earlier levels, and at which selection of reading material is no longer guided primarily by the occurrence of different expressional variants.

Estimating Difficulty of Reading Materials

The identification of elements of difficulty was followed by the development of a practical technique by means of which such elements can be used in estimating the difficulty of materials in which they occur. Elements of greatest possible usefulness were sought for this purpose. Since results of the study are intended primarily to benefit adults of limited reading ability, it seemed desirable to utilize elements whose difficulty had been defined for such readers. It was evident that twenty-four elements significantly related to difficulty would be too many for practical use. The number was reduced by defining useful elements as those that correlated closely with the criterion, that correlated relatively little among themselves, that are readily recognizable, and

that together give a convenient and adequate representation of indices of difficulty. Table X shows the intercorrelation of the significant elements of difficulty, by means of which the first two requisites to selection were met. Eight elements seemed to meet the definition of usefulness. They were: number of different hard words, number of easy words, per cent of monosyllables, number of first, second, and third person pronouns, average sentence-length in words, per cent of different words, number of prepositional phrases, and number of simple sentences.

Estimating relative difficulty. - The relative difficulty of selections may be estimated by merely counting the most significant indices of difficulty in their content, and then ranking the selections. If all eight elements are counted, the selection ranked as least difficult will contain the smallest number of different hard words, the most easy words, the highest per cent of monosyllables, the largest number of first, second, and third person pronouns, the shortest sentences, the lowest per cent of different words, the fewest prepositional phrases, and the most simple sentences. The number of elements used in such a procedure depends largely upon the amount of time that can be used in making the counts. That this method of estimating difficulty is fairly reliable was shown by the rank correlation computed between the absolute difficulty of test items in Forms 1 and 2 obtained by actual testing, and the relative difficulty of the items indicated by the ranks assigned to them in constructing the tests. The correlations were:

For Form 1, $p = .535 \pm .103$

For Form 2, $p = .663 \pm .081$

Estimating difficulty of specific materials. - Since all elements of difficulty contribute to the difficulty of a selection in which they occur, it seemed reasonable that a combination of the most useful elements into a single instrument of prediction would give the best estimate of difficulty. The eight elements were accordingly combined in a regression equation, in which each was assigned its relative weight, as follows:

$$\begin{aligned} X_1 = & -.006566 X_2 + .001268 X_3 + .004064 X_4 + .007545 X_5 \\ & -.02342 X_6 - .03371 X_7 - .01455 X_8 - .01015 X_9 + 3.408 \\ & \pm .2941. \end{aligned}$$

The meanings of $X_1 \dots X_9$ are:

X_1 = the average reading score which will probably be made by a group of adults of limited reading ability on a given passage of general reading material.

X_2 = the number of different hard words found in a passage of

TABLE X

CORRELATIONS OF SIGNIFICANT ELEMENTS OF DIFFICULTY WITH
AVERAGE READING SCORE AND WITH EACH OTHER

	Reading Score	Significant Elements of Difficulty*										
		1	2	3	4	5	6	7	8	9	10	11
Reading Score	1.000	-.550	-.543	.529	.518	.488	-.457	.450	-.449	-.441	-.428	.419
Ele- ments*												
1			.876	-.894	-.847	-.833	.660	-.713	.485	.424	.801	-.463
2				-.963	-.890	-.819	.558	-.539	.390	.598	.740	-.362
3					.952	.836	-.573	.612	-.398	-.559	-.743	.421
4						.804	-.518	.593	-.354	-.585	-.702	.439
5							-.657	.593	-.446	-.399	-.861	.423
6								-.659	.916	.171	.687	-.786
7									-.640	-.190	-.578	.702
8										.076	.462	-.845
9											.349	-.088
10												-.434
11												. . .
12												. . .
13												. . .
14												. . .
15												. . .
16												. . .
17												. . .
18												. . .
19												. . .
20												. . .
21												. . .
22												. . .
23												. . .
24												. . .
Mean	0.87	9.83	26.12	74.06	83.81	70.92	27.31	9.54	19.24	64.75	10.67	6.50
σ	0.57	6.21	9.28	8.07	10.31	6.99	11.85	5.96	7.09	5.06	6.26	2.61

*Significant elements of difficulty:

1. Number of words not known to 90 per cent of sixth-grade pupils.
2. Number of different hard words.
3. Per cent of easy words.
4. Number of easy words.
5. Per cent of monosyllables.
6. Average sentence length in syllables.
7. Number of first, second, and third person pronouns.
8. Average sentence length in words.
9. Per cent of different words.
10. Per cent of polysyllables.
11. Number of explicit sentences.
12. Per cent of different words not known to 90 per cent of sixth-grade pupils.

TABLE X (Continued)

Significant Elements of Difficulty*												
12	13	14	15	16	17	18	19	20	21	22	23	24
-.409	-.407	-.406	-.377	.371	-.370	-.367	.307	.300	.298	-.285	-.279	-.271
.877	.631	.414	.413	-.437	.506	.404	-.454	-.179	-.240	.094	.565	.458
.668	.604	.294	.442	-.315	.516	.603	-.380	-.275	-.187	.088	.612	.433
-.719	-.596	-.287	-.387	.375	-.509	-.507	.395	.262	.223	-.051	-.593	-.431
-.708	-.536	-.276	-.297	.411	-.467	-.408	.431	.241	.209	-.073	-.521	-.395
-.685	-.580	-.404	-.283	.411	-.499	-.372	.400	.136	.247	-.206	-.520	-.467
.555	.810	.809	.268	-.752	.485	.188	-.402	-.171	-.692	.167	.601	.440
-.640	-.632	-.523	-.304	.696	-.469	-.143	.556	.087	.427	-.150	-.510	-.455
.436	.794	.851	.173	-.810	.439	.079	-.390	-.135	-.775	.168	.550	.405
.378	.218	.085	.054	-.041	.108	.896	-.327	-.203	-.043	.240	.225	.035
.666	.577	.452	.350	-.414	.429	.354	-.340	-.115	-.309	.156	.513	.402
-.459	-.703	-.680	-.112	.977	-.511	.015	.319	.000	.824	-.092	-.523	-.531
. . .	.478	.380	.284	-.439	.405	.310	-.340	-.039	-.268	.075	.388	.406
. . .	. . .	.552	.331	-.647	.446	.267	-.426	-.003	-.550	.129	.899	.463
. . .	. . .	. . .	.048	-.666	.281	.060	-.334	-.173	.721	.276	.195	.227
. . .	. . .	. . .	. . .	-.088	.443	.128	-.048	-.335	.034	.068	.389	.334
. . .	. . .	. . .	. . .	. . .	-.501	.080	.297	-.026	.777	-.086	-.468	-.530
. . .	. . .	. . .	. . .	. . .	. . .	.069	-.199	-.283	-.266	.325	.411	.933
. . .	. . .	. . .	. . .	. . .	. . .	. . .	-.241	-.140	-.023	.204	.299	.019
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.176	.155	-.243	-.342	-.141
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.059	-.115	.060	.082
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	-.119	-.317	-.255
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.043	.294
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	.450
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	1.000
33.56	49.00	12.04	0.81	6.75	11.38	73.79	6.42	13.21	2.92	4.08	37.21	13.21
17.57	19.23	8.12	1.52	3.09	3.58	5.72	4.69	3.45	2.00	1.73	16.42	3.45

*Significant elements of difficulty (continued):

13. Maximum syllabic sentence length.
14. Minimum syllabic sentence length.
15. Number of figures of speech.
16. Number of sentences per paragraph.
17. Number of prepositional phrases.
18. Number of different words.
19. Number of third person pronouns.
20. Number of infinitive phrases.
21. Number of simple sentences.
22. Number of i words.
23. Range of syllabic sentence length.
24. Number of infinitive and prepositional phrases.

the size used in this study, i.e., about 100 words in length.

X_3 = the number of easy words found in the passage.

X_4 = the per cent of monosyllables used in the passage.

X_5 = the number of personal pronouns in the passage.

X_6 = the average sentence length in words used in the passage.

X_7 = the per cent of different words used in the passage.

X_8 = the number of prepositional phrases found in the passage.

X_9 = the number of simple sentences used in the passage.

To illustrate the use of this equation in predicting the average reading score of a selection, take for example a selection of reading material that contains measures for the independent variables as follows: 11 different hard words, 82 easy words, 76 per cent of monosyllables, 16 personal pronouns, an average sentence-length of 12.04 words, 36.50 per cent of different words, 7 prepositional phrases, and 4 simple sentences. The most probable average score that will be obtained by a group of readers of limited ability is obtained as follows:

$$\begin{aligned} X_1 = & -(.006566 \times 11) + (.001268 \times 82) + (.004064 \times 76) \\ & + (.007545 \times 16) - (.02342 \times 12.04) - (.03371 \times 36.50) \\ & - (.01455 \times 7) - (.01015 \times 4) + 3.408 = 2.2. \end{aligned}$$

Since the easiest item in the tests had an average score of 2.1 for "poorest" readers it is evident that the selection will be easier for similar readers than any item used in the tests.

In interpreting this average score, one may question what it means for individual readers, that is, how well they may be expected to comprehend a selection with a predicted average score of 2.2. An examination of the individual scores on test items having average scores between 1.75 and 2.1 (the highest obtained score) revealed the fact that for these items 58.6 per cent of the "poorest" readers showed a fair comprehension of the selections, by a score of 2 or better. A correlation of .898 was found between the per cent who can be expected to comprehend the selection and the average comprehension score which the group can be expected to make when tested for understanding of the selection. It is probable, therefore, that the majority of the group can be expected to comprehend a selection with a predicted average score of 2.2.

Reliability of prediction.- The precision of prediction by the regression equation just described depends partly on the size of the correlation between the actual difficulty and the predicted difficulty, expressed in terms of multiple r , designated R . It depends partly, also, on the size of the error of prediction. These values were:

$$R = .6446 \pm .084$$

$$PE(\text{Est: } X_1) = .2941.$$

It is evident that a combination of the eight most useful elements yielded a higher correlation with difficulty than did any single element shown in Table IX. The coefficient of alienation corresponding to $R = .6446$ was found to be .7645. This means that the estimate of average scores determined by the regression equation will on the average be in error about .76 as much as if the error resulted from pure guesses, or that the errors will be .24 smaller than those in pure guesses. The probable error of estimate shows that the chances are about even that the absolute average score will not vary from the predicted score more than $\pm .29$. For the selection used in the illustration, therefore, the absolute score will not be less than 1.91 nor more than 2.49.

Prediction by fewer variables.- In view of the considerable amount of labor involved in counting the occurrence of eight elements in a selection, effort was directed toward finding a combination of elements which could be counted expeditiously, and which would give a prediction approximately as good as that obtained by the larger number of variables. Because of the low weights attached to three elements: number of easy words, per cent of monosyllables, and number of simple sentences, it was evident that they probably did not increase predictive accuracy. When they were omitted from calculation, the new equation became:

$$X_1 = -.01029 X_2 + .009012 X_5 - .02094 X_6 - .03313 X_7 - .01485 X_8 + 3.774.$$

$$R = .6435 \pm .085$$

$$PE(\text{Est. } X_1) = .2945$$

Since multiple r and the probable error of estimate are approximately the same as the measures obtained for the equation using all eight elements, it is clear that quite as reliable results can be obtained by predicting difficulty of reading materials from the following elements: number of different hard words, number of personal pronouns, average sentence-length in words, per cent of different words, and number of prepositional phrases, as by including also number of easy words, per cent of monosyllables, and number of simple sentences.

Smaller combinations of the eight most useful elements were tried and their multiple correlations computed. Table XI gives a series of equations, combining four elements, which predict average reading score with a relatively high degree of reliability. The symbols have the same meaning as in the first equation given on page 26.

It must be emphasized that passages whose difficulty is predicted should be of similar size to those from which the regression equations were developed, that is, about 100 words in length. In predicting the difficulty of long selections, several

TABLE XI

INSTRUMENTS OF PREDICTION, COMBINING FOUR ELEMENTS OF DIFFICULTY

Variables	Formula	R	PE(Est. X_1)
1.2678	$X_1 = -.01229 X_2 - .02448 X_6$ $-.03245 X_7 - .01623 X_8 + 3.952$	.6402	.2956
1.2567	$X_1 = -.01333 X_2 + .01007 X_5$ $-.02224 X_6 - .03054 X_7 + 3.532$	.6393	.2958
1.3678	$X_1 = -.009709 X_3 - .02510 X_6$ $-.03410 X_7 - .01882 X_8 + 2.965$	.6387	.2960
1.4678	$X_1 = -.01279 X_4 - .02405 X_6$ $-.03866 X_7 - .01962 X_8 + 3.156$	.6385	.2961
1.2467	$X_1 = -.01156 X_2 + .006938 X_4$ $-.02552 X_6 - .03055 X_7 + 3.153$	.6368	.2967
1.2679	$X_1 = -.01540 X_2 - .02924 X_6$ $-.02999 X_7 - .01188 X_9 + 3.816$	.6355	.2970
1.4567	$X_1 = .01350 X_4 + .01057 X_5$ $-.02247 X_6 - .03758 X_7 + 2.682$	.6347	.2973
1.3467	$X_1 = .007945 X_3 + .009218 X_4$ $-.02622 X_6 - .03245 X_7 + 2.160$	.6345	.2974
1.3567	$X_1 = .01027 X_3 + .008694 X_5$ $-.02437 X_6 - .03298 X_7 + 2.535$	.6340	.2975

hundred word passages may be selected at random and examined for the presence of the elements of prediction. The average occurrence should then be substituted in the regression equation.

It should be remembered also that the equation predicts the difficulty of general material for readers of limited ability. That the "poorest" readers used in the study were a fairly adequate sampling of the population they were taken to represent was shown in two ways. First, the reliability coefficient obtained by splitting the group into random halves and then correlating their average scores on the test items was $.8221 \pm .0316$. Prediction by the Spearman-Brown prophecy formula gave a probable correlation of $.9024 \pm .0181$ between two groups, each of which is similar in size to the entire group of "poorest" readers. When the same procedure was repeated and a second index of reliability was computed, a correlation of $.9254 \pm .0140$ was found between the half-groups, and a probable correlation of $.9613 \pm .0074$ by application of the Spearman-Brown formula.

By the second method, the reliability of "poorest" readers

was obtained by measuring the "goodness" of the prediction obtained from their average reading scores. This was done by giving the Adult Reading Tests to eight new groups of adults of limited ability, and then correlating their absolute scores with the predicted scores based on the first population. A coefficient of $.720 \pm .075$ indicated that the regression equation based on the first population is valid for use with a second similar population.

Practical application of agencies of prediction.- It will be recalled that the present study originated in the hypothesis that reading materials can be more fittingly adapted to the ability of adult readers if a technique were devised by means of which difficulty of specific materials can be accurately estimated. Such a technique has now been developed, and its use illustrated. Any book of general content may be analyzed and its difficulty predicted, as far as difficulty can be measured by elements in written expression. By the use of one of the regression equations presented in this report, it is possible for librarians and others interested in adult reading to develop a workable list that will be of practical value in recommending books to adult readers.¹ For example, if a group of adults reads A Lantern in Her Hand, (predicted score, 0.86) with fair comprehension and enjoyment, it would be futile to recommend Jean-Christophe (predicted score, 0.23) and expect the same readers to understand it. It would be quite as futile to recommend The Country Life Reader, Book I (predicted score, 1.83), and expect them to enjoy it.

The technique devised in this study suggests a means by which writers and publishers can prepare simplified material for adults of limited reading ability. The procedure is (1) to sample material by taking a series of randomly selected passages approximately 100 words in length; (2) to analyze each passage for elements contained in the regression equation adapted for use; (3) to substitute the average occurrence of each element in the equation and solve for X_1 ; and (4) simplify the material if X_1 is low by reducing the occurrence of elements of difficulty and by increasing the occurrence of elements of ease. Tentative standards for this purpose are presented in the complete report of the investigation.

¹Such a list has been begun by the writer who predicted the difficulty of sixty books by use of the equation for variables 1.25678. The difficulty of these books, expressed in terms of the average reading score is given in the complete unpublished form of this dissertation which is filed in the Library of the University of Chicago.

Objectives of Future Investigation

The fruitfulness of an investigation lies not alone in its findings, but in its suggestion of greater accuracy and completeness by an extension and refinement of the problem in the same or in a new direction. During the course of this study, problems have arisen that could not be solved due to the limitations, imposed by time and purpose. Important among them are those pertaining to a more accurate and comprehensive means of estimating difficulty. They may be stated as follows:

1. How are variants of expression related to difficulty in different kinds of content:- history, travel, science, etc?
2. Are materials of the same difficulty when read for different purposes?
3. If a book makes more than one kind of appeal, is a reader's satisfactory response to one, an indication of the adequacy of his response to another?
4. To what extent is the reader's interest in a selection related to his ability to read it satisfactorily?
5. Is the actual difficulty of a selection related to a reader's impression that the selection is easy or hard?
6. To what extent does a reader's familiarity with a subject affect the influence of elements of difficulty?
7. Are elements of expression related to difficulty in the same degree for readers of different racial and linguistic background, and from different cultural environment?
8. How much weight can be given to elements of expression in determining the total difficulty of a selection?

After additional study has demonstrated that certain elements of expression are generally correlated with difficulty for all types of material when read for several purposes, it would then be possible to experiment with writing on a variety of topics in a manner that would be readily comprehended by adults of limited ability. The affect of utilizing different elements of expression in varying amounts could be determined by tests and interviews with persons who read the materials. It is important that the present study be supplemented by further investigation and experimentation if the reading ventures of readers of limited ability are to develop into permanent reading habits.

BIBLIOGRAPHY

Articles

- Dale, Edgar. "A Comparison of Two Word Lists," Educational Research Bulletin, X (December 9, 1931), 484-89.
- Edgerton, Harold A., and Toops, Herbert A. "A Formula for Finding the Average Intercorrelation Coefficient of Unranked Raw Scores Without Solving Any of the Individual Correlations," Journal of Educational Psychology, XIX (February, 1928), No. 2, pp. 131-38.
- Gray, William S. "The Importance of Intelligent Silent Reading," Elementary School Journal, XXIV (January, 1924), 348-356.
- L., H. W. "Million Books and Best Books," Golden Book, VIII (August, 1926), 382-83.
- Lewerenz, Alfred S. "Measurement of the Difficulty of Reading Material," Educational Research Bulletin, Los Angeles City Schools, XIII (March, 1929), 11-16.
- Montgomery, Wilda Lea. "An Investigation of the Uses of Recreatory Reading," University of Pittsburgh School of Education Journal, IV (November-December, 1928), 90-91.

Books

- Ayer, N. W., and Sons. Directory of Newspapers and Periodicals. Philadelphia: N. W. Ayer and Sons, 1931. Pp. 1386.
- Dickerson, L. L. Libraries and Adult Education. Report of a Study Made by the American Library Association. Chicago: American Library Association, 1926. Pp. 284.
- Gray, William S. and Monroe, Ruth. The Reading Interests and Habits of Adults. New York: Macmillan Co., 1929. Pp. xiii + 306.
- Holzinger, Karl J. Statistical Methods for Students in Education. Boston: Ginn & Co., 1928. Pp. viii + 372.
- Twenty-fourth Yearbook of the National Society for the Study of Education, Part I. Bloomington, Illinois: Public School Publishing Co., 1925. Pp. x + 336.

Unpublished Material

- Bear, Liata V. "The Length of Word as an Index of Difficulty in Silent Reading," Unpublished Master's thesis, Dept. of Education, University of Chicago, 1927. Pp. 77.
- Dale, Edgar. "Familiarity of 8000 Common Words to Pupils in the Fourth, Sixth, and Eighth Grades." Unpublished study, Bureau of Educational Research, Ohio State University. Pp. 84.

Dale, Edgar, and Tyler, Ralph W. "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability." Unpublished investigation, Bureau of Educational Research, Ohio State University, Columbus, Ohio, 1931. Pp. 30.

Leary, Bernice E. "An Analytical Study of Certain Elements in the Reading Material of Adult Magazines." Unpublished Master's thesis, Dept. of Education, University of Chicago, 1931. Pp. 122.

Parsons, Rhey B. "A Study of Adult Reading." Unpublished Master's thesis, Dept. of Education, University of Chicago, 1923. Pp. 124.

